Advanced Research on Information Systems Security

ISSN: 2795-4609 | ISSN: 2795-4560
Print & Online

Steganography and Computer Forensics - the art of hiding information: a systematic review

Cláudia Fernandes*

ISLA – Polytechnic Institute of Management and Technology, Vila Nova de Gaia, Portugal
Email: claudiasof.fer99@gmail.com

Abstract

This paper focuses on the study of steganoprahy as an anti-forensic method. The purpose of steganography techniques is to hide information from individuals unrelated to its content. Through a systematic review, the objectives of this article are: (I) explore and investigate the importance of steganography in computer forensics; (II) understand and analyze the methodology used to hide information in a file and (III) understand and analyze the methodology used to extract the information. The results summarize the previous research on this topics and we conclude the article with a warning to develpop more studies regarding this thematic as well as and make the security forces aware of this type of digital evidence and, in the same way that the methodology is being developed, the steganalysis must also seek, at least, to reach the same technological level

Keywords: steganography; steganalysis; computer forensics

---

* Corresponding author. Email address: claudiasof.fer99@gmail.com

1. Introduction

Stegano from the Greek word “steganos” means concealed, covered, or even protected, while graphy comes from the Greek word “graphein” that means writing. The art of hiding or disguising information within a message is called steganography [1]. The purpose of steganography techniques is to hide information from people foreign to its content [2].

The first report that history had found regarding steganography was during the 15th century, where messages were hidden and covered in wax. Writing with invisible ink –lemon juice – was also used in order to hide messages in objects that seem normal at first glance. In order to reveal the message, flame or heat were needed. Other examples also involved the usage of animal carcass and even the human body as the stego object [3]. However, we need to keep in mind that, due to the carrier used, we could encounter drawbacks regarding the size of the message; its reusability and even the reveal method.

We can consider that the most important characteristic of stenography is its statical undetectability [4], hence, more complex methods are preferable as they are more difficult to decipher. Steganography is focused in hiding the information rather than securing it, while cryptography focus on the information security. Compared to encryption, its advantage is also the fact that the message format does not attract any suspicion since it cannot be seen with the naked eye [2].

Even though there are different methodologies of steganography used throughout the centuries, in this article, however, the focus of our study is digital steganography and their use as an anti-forensic measure.

In the last few years, there has been a growing dependence on technology in our daily lives, from entertainment and leisure to our professional and academic life [5]. In the same way that technology is attractive and advantageous in terms of practicality and automation of information, it is also attractive for individuals who intend to exploit vulnerabilities in the law and regulations, hiding behind a screen and committing cybercrimes using their computer as a "weapon".

Digital files such as images, audio, or video [4] have specific structures and areas that are susceptible to modification in order to hide information, without the use of the files being significantly affected or a user who happens to come across these files easily realizing that there is information hidden in them [6]. There are three important factors in steganography, (I) the file format, (II) how much space we have to hide information and (III) the steganography method [7].

The detection of this type of hidden information currently tends to rely on programs that recognize common stenography patterns and even on the statistical analysis of files with hidden information allowing to measure the deviation from the normal pattern of files. However, certain more sophisticated techniques can prevent programs from locating the information, and, in this case, the detection of steganography techniques can be done by analyzing the properties of the files and their internal description. It is important to point out that even when steganographic content is detected, in more complex scenarios, this content is often encrypted so that it cannot be read directly without a cryptographic key [6].

Steganography can be used for several purposes, namely (I) digital watermarks - in order to prevent the individual's intellectual property from being stolen, copied and sold; (II) sharing passwords / sensitive information as a means of encryption; (III) as a means of communication between individuals of the same group (criminal or not) and (IV) to impress friends or family or for intertreatment. However, in addition to harmless uses, it is also used as an anti-forensic measure, in order to hide illegal information so that it cannot be intercepted by the

authorities [2].

In this article, we intend to develop and analyze steganography techniques as an anti-forensic measure. Through a systematic review, the objectives of this article are:

1) Explore and investigate the importance of steganography in computer forensics.

2) Understand and analyze the methodology used to hide information in a file.

3) Understand and analyze the methodology used to extract the information.

2. Methodology of research

This article methodology focused on a systematic review of previous research and documentation [22], [23], [24], [25]. Following the PRISMA flowchart, the research method was outlined:

The first step was the identification of multiple articles that could potentially become part of the final list. For this research, Researchgate was the main source of information, however grey literature was also found in articles references and added to the total articles (n = 178). The main keywords for the search were:

Keyword	Operator
Steganography	AND
Forensic Computing	OR
Steganography tools	OR
Steganography methods	OR

The exclusion and inclusion criteria were also be defined. While the exclusion criteria was mostly related to the availability of the texts, only abstract available and articles that were unrelated to que question, the inclusion criteria was focused in articles that contained the information that helps to answer the research question. For this research, all articles that can't be found in its full text form were excluded, and the included articles had to focused on the research questions. All records were exported to a EndNote Library in order to remove and avoid duplication of the same article.

The reviewing process takes three phases: the first one, analysing the titles, a second one analysing abstracts and the latter, analysing the full text as a whole. After that, the author analyzed which articles are the best and more suitable for the review. This is the screening phase in the PRISMA flowchart [8].

After evaluation of the articles, they were extracted and all date was synthesised and 21 articles were included in the research.

3. **Screening and Eligibility of Results**

We compiled and examined the abstracts of 178 papers to conduct this review of the State of The Art. Following review, we retained only the English. The remaining materials in this State of the Art were grouped into the following categories depending on their scope and intended use: Steganography techniques; Steganography detention techniques and the importance of steganography in computer forensics. In order to conduct a systematic review, the collected studies were processed recurring to PRISMA approach. Fig. 1 shows the revision method that was used.

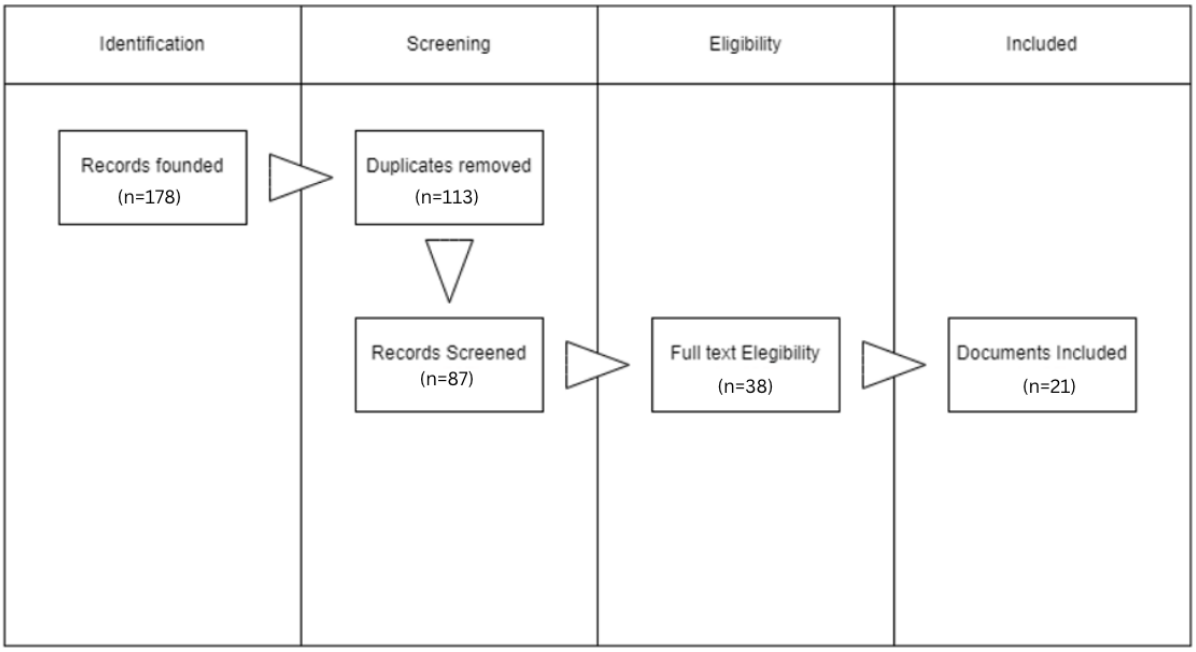


Figure 1: PRISM Diagram used to screen documents

4. Results

A total of 178 results were gathered from the publications database, and after duplicate removal and screening, 113 were chosen for eligibility analysis as follows: A total of 21 articles were chosen and used as references for this state of the Art after full-text eligibility was determined. The study's findings are depicted in Figure 1.

4.1 Importance of steganography in computer forensics

The importance of computers forensics arises for the need to use digital evidence in court in order to prosecute criminals and to achieve justice to the victims. Computers don't store only date regarding criminals that engage in cybercrime, and we need to keep in mind that many other criminals, maintain files with incriminating evidence on their computer [9]. A few examples are related to Al-Qaeda members that were hiding information regarding future plans of the organizations in frames of pornographic content [10], as well as the infamous cartel leader that was sending orders and information via Hello Kitty pictures [11] and even Russian spies [12].

As previously referred, with the growing dependence on technology in our daily lives and the growing vulnerability exploits, many internet and computer users found in steganography a way to protect sensitive information and even conversations. However, at the same time, criminals started to use this method in order to hide illegal information [13].

It is also important to note that, the biggest obstacle that digital forensic examiners face is regarding the analysis and identification of hidden data in digital media. Furthermore, there is also lack of specialized forensic investigators that are able to identify this type of evidence and their methods [14].

However, there are a few ways to detect data hiding in a suspect's software, including: (I) data hiding software; (II) cached files, (III) thumbnails, and other evidence may reside on the suspect computer [15].

4.2 Steganography techniques

It was around 1985 that modern steganography methods started to be implemented and by the 90's there were three forms stipulated: (I) pure steganography; key steganography (where a key is shared between the two communicating parties) and public key steganography (public key is used to encrypt and private key to decrypt the message) [12].

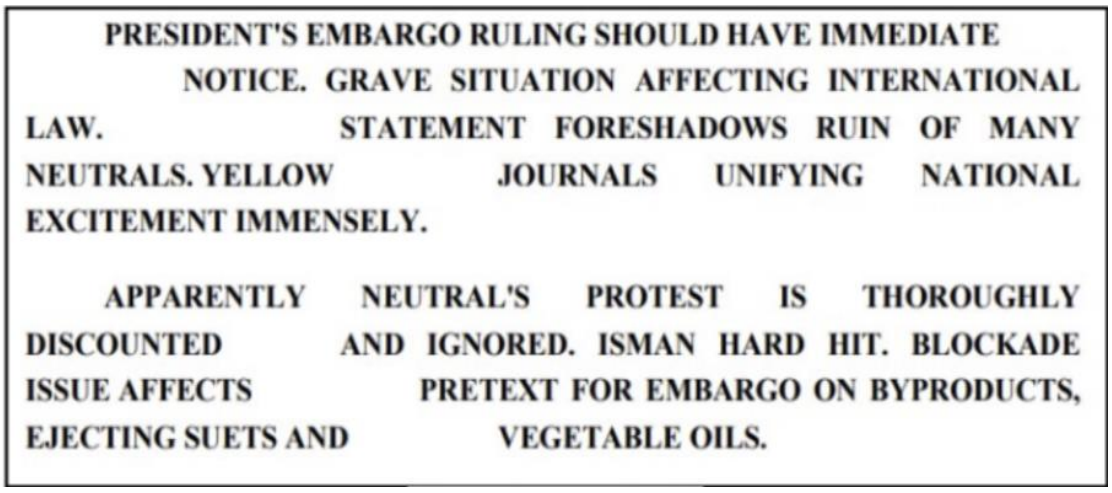
Injection techniques: The methodologies of injection techniques use existing information in the file or available space in it to inject new information in it. Although there is no change in the file content and we are just adding new information [15], the ability to hide information is very limited [2].

Substitution techniques: In this type of steganography, we replace redundant or unnecessary bits of the carrier with the bits from the secret message. [16] The most common form of substitution techniques involves modifying the Least Significant Bit (LSB) of one or more bytes within a file [15], in order to build a message. Since there are thousands of pixels in an image, we can store information without a common user noticing. This reasoning can be applied to audio, video and other types of digital files. Some more advanced techniques are Discrete Cosine Transformation (DCT), Fourier transformations, and Patchwork Method of Bender. For audio files, steganography is done by imperceptibly replacing short segments with reference phases that represent the hidden data (phase encoding).

Because the bits are being changed [17] there is, however, some degradation of the original file when using this type of techniques [2]. In image files, this bit contributes to the brightness of the image or noise within it, however in large files they cannot be easily detected [17]. Nonetheless, it is important to note that the risk of detection increases with the volume of hidden content [2].

File creation techniques: This technique makes it possible to generate a new file and, in a certain way, the message to become its own bearer [2]. Some examples include the methodology used by Germany during the World War I, using telegrams as a medium, at first, the message does not have any meaningful insight, however, by reading the first letter of each word, we would be able to uncover a hidden message [3] (Figure 2.). Although a pre-

existing file is not required, this method is inefficient, and the risk of detection is dependent on the context of the message [2].


PRESIDENT'S EMBARGO RULING SHOULD HAVE IMMEDIATE NOTICE. GRAVE SITUATION AFFECTING INTERNATIONAL LAW. STATEMENT FORESHADOWS RUIN OF MANY NEUTRALS. YELLOW JOURNALS UNIFYING NATIONAL EXCITEMENT IMMENSELY.

APPARENTLY NEUTRAL'S PROTEST IS THOROUGHLY DISCOUNTED AND IGNORED. ISMAN HARD HIT. BLOCKADE ISSUE AFFECTS PRETEXT FOR EMBARGO ON BYPRODUCTS, EJECTING SUETS AND VEGETABLE OILS.

Figure 2: Steganografy used in World War I. The message hidden is Pershing Sails from N.Y June 1 [3]

4.2 Detention techniques

Steganography can be detected via various tools, taking into account their three components: the sender, received and the carrier [3]. Steganalysis is the practice of detecting the use of steganography [18] and research states that, currently, there are two different types of approach in stenography identification: the first is the targeted steganalysis which to detects a particular algorithm; the second, is capable of detecting steganalysis without knowing the steganographic algorithms [19]. The method used depends on the hiding technique [15].

The steganalysis is also the process of analyzing small deviations from the expected patterns of a file so that hidden messages can be detected, recognizing recognize various types of signatures from steganography programs (i.e., common patterns that appear to appear when these tools have been used). The following types of steganananalysis can be distinguished: "stego only attack" (when we only have the document available); known cover attack; known message attack; chosen stego attack (the object and algorithm are available for analysis); chosen message attack (choose a normal message which is then converted) and known stego attack (where the shorthand message, algorithm and and the translated message are available) [2].

There are a few steganalysis tools available to the public, namely SecretLayer, Openpuff, Openstego, Steg, Steganography studio and GhostHost [13]. However, OpenPuff is considered to be the most complete option as it provides support to different type of files, and it is also open source [20].

There is also the passive attack that one that observes the communications, the sender and the received as well as the carrier – after getting the details, the analyst is able to uncover the message [3]. And also, the statistical analysis of files with hidden information allowing to measure the deviation from the normal file pattern known and “clean” (that is, without steganographic information) [2].

Nevertheless, we need to keep in mind that with the increasing of steganographic software’s, the steganalysis

methods need to keep up with the demand [19].

5. Discussing and conclusions

Digital steganography has a multitude of applications, both legal and illegal, and is considered an advantageous technique that is relatively easy to access for the common user. It is important to note that, while it has multiple advantages (e.g., exposing plagiarism), it is also a technique widely used as an anti-forensic measure.

The lack of expertise on this field and constant evolution of technology can be overwhelming for a forensic investigator. They may look for multiple versions of the same image, search metadata and registries, use histograms to be able to identify altered files and still, no certain response is found. The first step is to learn what steganography techniques were used and that steg-data is present. The second step with the knowledge acquired is to identify the software that was able to create the suspected data. The third step is trying to determine the key used to conceal the data, either by brute force or other kind of cryptographic attacks [21].

In this way, it is crucial to study and make the security forces aware of this type of digital evidence and, in the same way that the methodology is being developed, the steganalysis must also seek, at least, to reach the same technological level.

References

[1] Raphael J, Sundaram V. Cryptography and Steganography – A Survey. Int J Comput Tech Appl 2011;2(3):626-30.
[2] Warkentin, Merrill; Bekkering, Ernst; and Schmidt, Mark B. (2008) "Steganography: Forensic, Security, and Legal Issues," Journal of Digital Forensics, Security and Law: Vol. 3 : No. 2 , Article 2.
[3] N. Alabdali and S. Alzahrani, 'An Overview of Steganography through History', vol. 5, no. 2, p. 4.
[4] Bierbrauer, Jürgen, and Jessica Fridrich. "Constructing Good Covering Codes for Applications in Steganography." Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2008, pp. 1–22. https://doi.org/10.1007/978-3-540-69019-1_1.
[5] Kumar, V.D. Ethical Hacking and Penetration Testing Strategies. 2014
[6] Angus Marshall, Wiley-Blackwell. Digital Forensics, Digital Evidence in Criminal Investigation". 2008.
[7] Liu, Bo, et al. "Thwarting Audio Steganography Attacks in Cloud Storage Systems." 2011 International Conference on Cloud and Service Computing, IEEE, Dec. 2011, https://doi.org/10.1109/csc.2011.6138530
[8] Tawfik, Gehad Mohamed, et al. "A Step by Step Guide for Conducting a Systematic Review and Meta-analysis With Simulation Data." Tropical Medicine and Health, vol. 47, no. 1, Springer Science and Business Media LLC, Aug. 2019, https://doi.org/10.1186/s41182-019-0165-6.
[9] Kamble, Dhwaniket Ramesh, and Nilakshi Jain. "DIGITAL FORENSIC TOOLS: A COMPARATIVE APPROACH." International Journal of Advance Research in Science and Engineering IJARSE, Vol. No.4, Issue No.02, February 2015, Jan. 2015.
[10] Gallagher, Sean. "Steganography: How al-Qaeda Hid Secret Documents in a Porn Video." Ars Technica, 2 May 2012, arstechnica.com/information-technology/2012/05/steganography-how-al-qaeda-hid-secret-documents-in-a-porn-video.
[11] Folha De S.Paulo - Para Agência Dos EUA, Abadía Traficou No Brasil - 10/03/2008." Folha De S. Paulo, www1.folha.uol.com.br/fsp/cotidian/ff1003200801.htm.
[12] Stokes, Jon. "How Even the Dumbest Russian Spies Can Outwit the NSA." Ars Technica, 5 July 2010, arstechnica.com/tech-policy/2010/07/how-even-the-dumbest-russian-spies-outwit-the-nsa.
[13] Ashok J, Raju Y, Munishankaraiah S, Srinivas K. Steganography: An Overview. Int J Eng Sci Technol 2010;2(10):5985-92.
[14] Alqahtani, Jawaher, et al. "Steganalysis Algorithm for PNG Images Based on Fuzzy Logic Technique." International Journal of Network Security &Amp; Its Applications, vol. 8, no. 6, Academy and Industry

Research Collaboration Center (AIRCC), Nov. 2016, pp. 01–15. https://doi.org/10.5121/ijnsa.2016.8501.

[15] Raggo, Michael, and Chet Hosmer. Data Hiding: Exposing Concealed Data in Multimedia, Operating Systems, Mobile Devices and Network Protocols. 1st ed., Syngress, 2012.

[16] Kipper, Gregory. Investigator's Guide to Steganography. 1st ed., Auerbach Publications, 2003.

[17] Yugala K. Steganography. Int J Eng Trends Technol 2013;4(5):1629-35.

[18] Tariq Jamil, "Steganography, the art of hidden information in plain sight", IEEE Potentials, Feb/March 1999, ISSN 0278-6648

[19] Gong, Chen, et al. "Detecting Fingerprints of Audio Steganography Software." Forensic Science International: Reports, vol. 2, Elsevier BV, Dec. 2020, p. 100075. https://doi.org/10.1016/j.fsir.2020.100075.

[20] Karadogan, Ismail & Das, Resul & Of,. An Examination on Information Hiding Tools for Steganography. International Journal of Information Security. 3. 200-208. 2014.

[21] Kent, K., et al. "Guide to Integrating Forensic Techniques Into Incident Response." National Institute of Standards and Technology, National Institute of Standards and Technology, 2006, https://doi.org/10.6028/nist.sp.800-86.

[22] Duarte, N., Coelho, N., Guarda, T. (2021). Social Engineering: The Art of Attacks. In: Guarda, T., Portela, F., Santos, M.F. (eds) Advanced Research in Technologies, Information, Innovation and Sustainability. ARTIIS 2021. Communications in Computer and Information Science, vol 1485. Springer, Cham. https://doi.org/10.1007/978-3-030-90241-4_36

[23] F. Alves, N. Mateus-Coelho and M. Cruz-Cunha, "ChevroCrypto – Security & Cryptography Broker," 2022 10th International Symposium on Digital Forensics and Security (ISDFS), Istanbul, Turkey, 2022, pp. 1-5, doi: 10.1109/ISDFS55398.2022.9800797.

[24] N. Mateus-Coelho and M. Cruz-Cunha, "Serverless Service Architectures and Security Minimals," 2022 10th International Symposium on Digital Forensics and Security (ISDFS), Istanbul, Turkey, 2022, pp. 1-6, doi: 10.1109/ISDFS55398.2022.9800779.

[25] Nuno Mateus-Coelho, A New Methodology for the Development of Secure and Paranoid Operating Systems, Procedia Computer Science, Volume 181, 2021, Pages 1207-1215, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2021.01.318.